J|A|C|S
COMMUNICATIONS

Selective 2′-Hydroxyl Acylation Analyzed by Protection from Exoribonuclease

Kady-Ann Steen,[†] Arun Malhotra,[‡] and Kevin M. Weeks*[,†]

Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27599-3290 and Department of Biochemistry and Molecular Biology, University of Miami, Miller School of Medicine, Miami, Florida 33101

Received May 3, 2010; E-mail: weeks@unc.edu

Abstract: Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) is a powerful approach for characterizing RNA structure and dynamics at single-nucleotide resolution. However, SHAPE technology is limited, sometimes severely, because primer extension detection obscures structural information for ~15 nts at the 5′ end and 40−60 nts at the 3′ end of the RNA. Moreover, detection by primer extension is more complex than the actual structure-selective chemical interrogation step. Here we quantify covalent adducts in RNA directly by adduct-inhibited exoribonuclease degradation. RNA 2′-*O*-adducts block processivity of a 3′→5′ exoribonuclease, RNase R, to produce fragments that terminate three nucleotides 3′ of the modification site. We analyzed the structure of the native thiamine pyrophosphate (TPP) riboswitch aptamer domain and identified large changes in local nucleotide dynamics and global RNA structure upon ligand binding. In addition to numerous changes that can be attributed to ligand recognition, we identify a single nucleotide bulge register shift, distant from the binding site, that stabilizes the ligand-bound structure. Selective 2′-hydroxyl acylation analyzed by protection from exoribonuclease (RNase-detected SHAPE) should prove broadly useful for facile structural analysis of small noncoding RNAs and for RNAs that have functionally critical structures at their 5′ and 3′ ends.

The three-dimensional structures of RNAs play direct roles in gene expression and regulation. The sizes of important regulatory elements vary enormously from large catalytic and structural RNAs[1] to the small, compact structures of microRNA precursors, tRNAs, and riboswitches.[2] The function of each of these motifs is dependent on the specific local structures and dynamics that characterize each nucleotide. The local structural environment at most nucleotides in large RNAs can be probed using selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE).[3] In a SHAPE experiment, an RNA is allowed to react with an acylating agent that is selective for the ribose 2′-hydroxyl group to form a 2′-*O*-adduct (Figure 1A). All four RNA nucleotides react similarly with the acylating reagent,[4] and the extent of reaction is quantitatively proportional to local nucleotide dynamics.[5] In principle, local nucleotide dynamics and flexibility can be assessed rapidly in a simple chemical interrogation step. Sites of 2′-*O*-adduct formation are then detected by primer extension.[6]

The use of primer extension to detect 2′-*O*-adducts represents a weakness of SHAPE technology, especially as applied to important short RNAs, and is also a shortcoming for many other probing approaches for analyzing RNA modification chemistries and structure. Using primer extension, RNA adducts are detected by

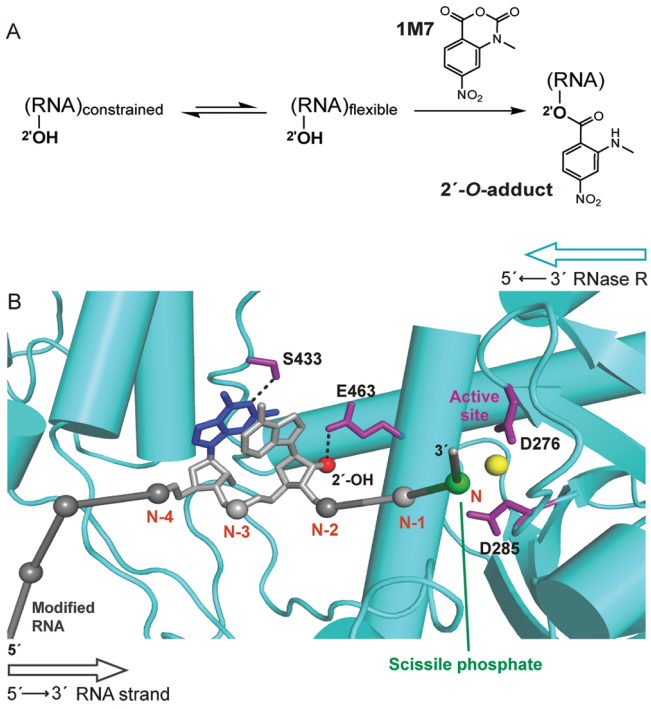

Figure 1. Schemes for (A) SHAPE chemistry and (B) RNase R recognition and degradation of RNA. 2′-Hydroxyl and nucleobase positions specifically recognized in the substrate binding channel are shown as a red sphere and in blue, respectively. Phosphate hydrolysis occurs at the green sphere; an active site Mg^{2+} is yellow, and other phosphate groups are gray.

annealing a labeled DNA primer to the 3′ end of a modified RNA and then a reverse transcriptase enzyme is used to extend these primers to the sites of modification. Chemical information is thus read out indirectly, as the lengths and frequency of a given cDNA product, instead of by direct analysis of the RNA fragment. No structural data are obtained for the 40−60 nucleotides at the 3′ end of the RNA due to overlap with the primer binding site and incomplete enzyme processivity in the early stages of extension. Structural data for 10−20 nucleotides at the 5′ end of the RNA are also obscured due to overlap with full-length extension products. These two features make it impossible to analyze the structures of biologically important short RNAs in their native forms. Finally, analysis of RNA modifications by primer extension is experimentally more complex than the actual SHAPE chemical probing step.

We have now developed a *direct* method for detecting covalent adducts in RNA, including the 2′-*O*-adducts created by SHAPE, based on adduct-selective protection from exoribonuclease degradation. The RNase R family of exoribonucleases processively and nonspecifically hydrolyzes RNA in the 3′→5′ direction to release 5′-nucleotide monophosphates.[7] We screened RNase R enzymes from three organisms and determined that the RNase R enzyme from *Mycoplasma genitalium* degrades structured RNAs, cannot

[†] University of North Carolina.
[‡] University of Miami.

10.1021/ja103781u

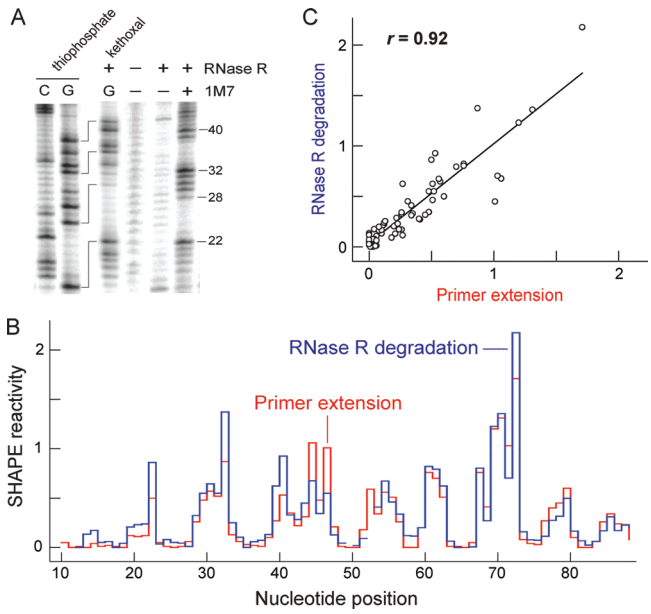

Figure 2. Quantitative detection of 2′-*O*-adducts by RNase R degradation. (A) Comparison of RNase R degradation of 5′-end labeled RNAs containing 1M7 adducts with C and G sequencing ladders, visualized by gel electrophoresis. Sequencing ladders were generated either by I_2-mediated cleavage of phosphorothioate-containing RNAs or as kethoxal-mediated stops to RNase R degradation. The kethoxal marker lane fragments (G) are one nucleotide shorter than those in the 1M7 lanes. (B) Histograms comparing absolute SHAPE reactivities as determined by RNase R degradation and primer extension. Nucleotides that could not be analyzed due to high background are indicated by breaks in the plot. (C) Correlation between SHAPE reactivities detected by RNase R degradation and primer extension. Pearson's linear *r*-value is shown.

proceed past 2′-*O*-methyl modifications,[8] and is readily inactivated by heat treatment (Supporting Information). We then evaluated whether RNase R degradation could be used to detect sites of 2′-*O*-adduct formation in the aptamer domain of the *Escherichia coli thiM* thiamine pyrophosphate (TPP) riboswitch RNA.[9] This small RNA (80 nts) contains many features common to structured RNAs, including canonical and noncanonical base pairing, local stacking, and long-range docking interactions.[9]

We modified a 5′-end labeled TPP riboswitch using 1-methyl-7-nitroisatoic anhydride (1M7)[10] in the absence of ligand (in 10 mM $MgCl_2$, 100 mM NaCl, pH 8.0; at 37 °C). The modified RNA was then subjected to degradation by RNase R (0.25 mM $MgCl_2$, 100 mM KCl, pH 8.0; 30 min, 50 °C), and the resulting end-labeled RNA fragments were resolved by gel electrophoresis (Figure 2A). Following the degradation step, the RNase R enzyme can be completely heat inactivated by incubation at elevated temperature (95 °C, 3 min; see Supporting Information). To facilitate comparison with conventional primer extension analysis of the sites of modification, these initial experiments were performed using a TPP riboswitch RNA containing non-native 5′ and 3′ "structure cassette" sequences[6] of 14 and 43 nucleotides, respectively.

The RNase R enzyme efficiently degraded the unmodified riboswitch RNA. In contrast, when the RNA was treated with 1M7 and then incubated with RNase R, a strong pattern of banding was observed [compare (−) and (+) 1M7 lanes, Figure 2A]. The lengths of these 1M7-modified fragments were determined by comparison with sequencing ladders generated by iodine-mediated cleavage of phosphorothioate-substituted RNA[11] and by RNase R degradation of kethoxal-modified RNA (sequencing lanes, Figure 2A). Unexpectedly, bands corresponding to guanosine residues in the two sequencing reactions were offset by five nucleotide positions on the gel. Phosphorothioate cleavage results in a 2′,3′-cyclic phosphate fragment[11] that is one nucleotide shorter than the guanosine-terminated fragment, when visualized using 5′-labeled RNA. Taking into account this offset for phosphorothioate cleavage, the net offset for the phosphorothioate- and kethoxal-mediated sequencing reactions is four nucleotides.

To determine if RNase R degradation, in fact, recapitulated RNA 1M7 reactivity and to understand the large offset in the sequencing reactions, we performed identical SHAPE reactions on the TPP riboswitch but analyzed the results by conventional primer extension. Absolute reactivities were calculated by subtracting the no-reagent intensities from the 1M7-modified intensities for reactions analyzed by both RNase R degradation and primer extension (blue and red, Figure 2B). Absolute reactivities detected by RNase R degradation were almost identical to those obtained by primer extension (Pearson's linear $r = 0.92$; Figure 2C). However, to superimpose the reactivity profiles, RNase R-detected bands in the 1M7 reaction and in kethoxal-mediated sequencing required 3- and 4-nucleotide offsets, respectively, to yield agreement with the primer extension-detected reactions.

To understand the chemical basis of these offsets, we created a homology model[12] for the *M. genitalium* RNase R enzyme based on the known structures of two close homologues, *E. coli* RNase II (2ix0, 2ix1) and *S. cerevisiae* Rrp44 (2vnu).[13] Consistent with prior work,[13] the single-stranded RNA threads through the substrate binding channel, ultimately occupying the active site where phosphate cleavage occurs (termed nucleotide N and shown with a green sphere, Figure 1B).

Critically, the enzyme also makes several contacts with RNA nucleotides 5′ of the active site. First, the RNase R enzyme forms a hydrogen bond between the 2′-OH group at nucleotide N-3 (red sphere, Figure 1B) and glutamic acid 463, consistent with studies[14] showing this residue is essential for RNA cleavage specificity. Second, serine 433 forms a hydrogen bond with the base-pairing face of nucleotide N-4 (blue nucleotide, Figure 1B). Thus, the RNase structure is consistent with a model in which a 1M7-mediated 2′-*O*-adduct or a kethoxal-mediated cyclic adduct at guanosine[15] causes RNase R degradation to stop either three or four nucleotides, respectively, 3′ of the site of modification (Figure 1B).

In sum, RNase R-mediated degradation of end-labeled RNA yields quantitative detection (Figure 2B, C) of covalent adducts at both the ribose 2′-OH position in the backbone and at the base-pairing face of guanosine. Sites of adduct formation can be readily assigned by noting that kethoxal-mediated sequencing bands are exactly 1 nt shorter than the corresponding 2′-*O*-adduct (Figure 1B). RNase R degradation thus provides a novel, direct, and efficient one-step approach for detecting covalent adducts in RNA. Selective 2′-hydroxyl acylation analyzed by protection from exoribonuclease (termed RNase-detected SHAPE) makes it possible to probe the single nucleotide resolution structure of native RNAs, including small noncoding RNAs such as the TPP riboswitch, without addition of extraneous 5′ and 3′ flanking sequences.

Crystallographic and biochemical analyses have yielded a wealth of information about the secondary and tertiary structures of the ligand-bound state of the TPP riboswitch.[9,16] However, a single nucleotide resolution structure of the ligand-free state and the changes in RNA dynamics that occur upon ligand binding are unknown. We therefore used RNase-detected SHAPE to analyze the structure of the *native* TPP riboswitch aptamer domain in the absence and presence of the TPP ligand. Significant differences in SHAPE reactivities for the free and ligand-bound states were observed (Figure 3). Absolute SHAPE reactivities were quantified

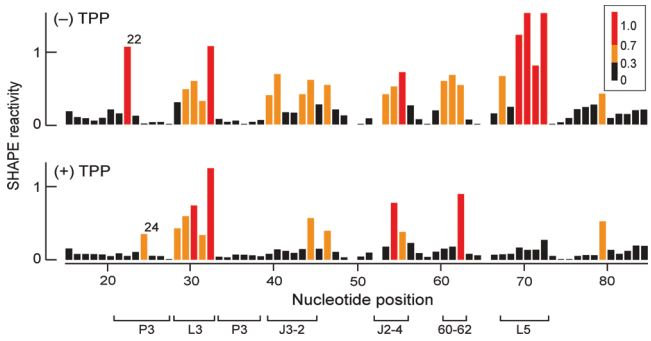

Figure 3. Detection of 2′-*O*-adducts by protection from exoribonuclease degradation in the *native* TPP riboswitch domain. Histograms show absolute SHAPE reactivities in the absence (top) and presence (bottom) of TPP. Structural landmarks that show large changes in reactivity, and thus local nucleotide dynamics, are highlighted below the *x*-axis.

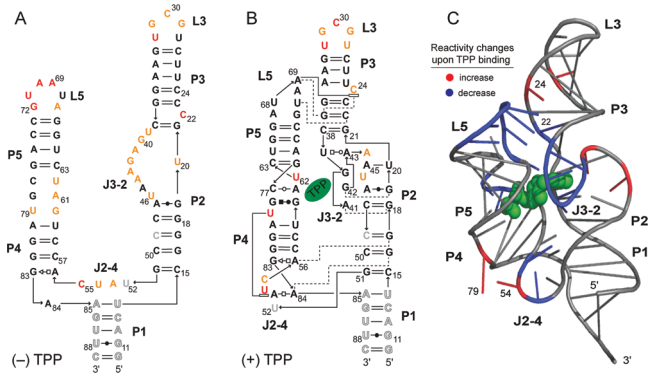

Figure 4. Absolute SHAPE reactivities determined by RNase R degradation superimposed on secondary structure models for the 80 nt TPP riboswitch in the (A) absence and (B) presence of TPP ligand. Nucleotides are colored by SHAPE reactivity using the scale shown in Figure 3. All nucleotides in the RNA are shown; the small number of positions for which no data were obtained are gray. (C) Increases (red) or decreases (blue) in SHAPE reactivity upon TPP ligand binding superimposed on the three-dimensional riboswitch structure.

for most positions in this small RNA and were used to create an experimentally supported secondary structure model[17] of the ligand-free state (Figure 4A). The ligand-free secondary structure is characterized by three significant changes relative to the ligand-bound state. First, all loops in the unliganded RNA were highly flexible. Second, with the exception of the G19-A47 and A56-G83 base pairs, noncanonical base pairs did not form stably in the ligand-free structure. Third, the P3 helix showed a significant register shift in the ligand-free relative to the ligand-bound state. In the TPP-bound state, nucleotides in P3 have SHAPE reactivities that are exactly consistent with their crystallographically visualized structure: C24 is reactive while all base-paired positions are unreactive (see C24 in Figure 4B). In contrast, in the ligand-free state, C24 is unreactive while C22 is reactive, indicating the latter nucleotide is the bulged nucleotide (Figure 4A). In sum, the SHAPE data emphasize that, in the ligand-free state, the TPP riboswitch consists of five well-formed and stable helices that are linked by highly dynamic single-stranded regions, suggestive of a structure with little or no stable tertiary interactions.

Upon addition of TPP, large changes in local nucleotide dynamics occur throughout the riboswitch RNA, especially in P3, J3-2, J2-4, L5, and nucleotides 60-62 (compare panels, Figure 3). Several regions that became more constrained in the ligand-bound state reflect direct interactions between the ligand (specifically with the TPP aminopyrimidine and pyrophosphate moieties) and the RNA. These elements include J3-2 and nucleotides 60-61 (Figures 4B,C and S1). A few nucleotides near the TPP binding site became more flexible in the ligand-bound state. For example, the SHAPE reactivity of nucleotide 62 increased, consistent with becoming extrahelical to form the pyrophosphate binding pocket for TPP (Figure 3B and S1).[9,16a]

Critically, regions that are spatially distant from the ligand binding site also showed large changes in SHAPE reactivity. The L5 loop was highly dynamic in the ligand-free state but became completely constrained in the ligand-bound state, consistent with L5 docking into the minor groove of P3 (Figure 3). The register shift in P3 results in the bulged nucleotide migrating from C22 to C24 and allows C24 to stack with A69 (Figure 4B). The L5-P3 interaction forms early during TPP binding,[16b] and the C24-A69 stacking interaction makes a significant contribution to the stability of the RNA−ligand complex.[16d] The large change in SHAPE reactivity observed in L5 nucleotides thus implies that the C22 to C24 bulge migration functions to direct riboswitch folding and to stabilize the structure of the ligand-bound state.

The J2-4 region also undergoes large conformational changes to form interactions that stabilize the three-way junction (Figures 3,4). A53 became less reactive in the ligand-bound state, consistent with formation of the A53-G84 noncanonical base pair that stabilizes the three-way junction.[9] The crystallographic data indicate that U54 and U79 form a stacking interaction in the ligand-bound state;[9] SHAPE data indicated that these nucleotides are highly flexible in both free and ligand-bound states. The flexibility in this stacking interaction suggests that while the interaction may occur transiently, it does not stabilize the ligand-bound state in the isolated aptamer domain.

Selective 2′-hydroxyl acylation analyzed by protection from exoribonuclease combines quantitative and robust structure-selective RNA acylation with a simple one-tube exoribonuclease degradation step for detecting sites of covalent modification in RNA. Degradation by *M. genitalium* RNase R was inhibited by 2′-*O*-adducts and adducts at the base-pairing face of guanosine due to specific, but distinct, interactions in the substrate binding channel of the enzyme (Figure 1B). RNase-detected SHAPE is thus likely to be broadly useful for identifying and quantifying many additional classes of chemical adducts in RNA. Using this technology, we generated single-nucleotide resolution structure models for both the ligand-free and ligand-bound states of the 80 nt native sequence TPP riboswitch aptamer domain (Figure 4A, B) and characterized a single nucleotide bulge shift from C22 in the ligand-free state to C24 in the ligand-bound state that is partly responsible for stabilizing the ligand-bound RNA structure.

Selective 2′-hydroxyl acylation analyzed by protection from exoribonuclease makes possible direct and facile single nucleotide resolution structural analysis of RNAs that were previously difficult or impossible to evaluate, including small RNAs in their native states and the 5′ and 3′ ends of important regulatory RNAs.

Acknowledgment. The work was supported by the U.S. National Science Foundation (MCB-0919666 to K.M.W.) and National Institutes of Health (GM69972 to A.M.).

Supporting Information Available: Experimental methods and one figure. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Holbrook, S. R. *Annu. Rev. Biophys.* **2008**, *37*, 445–464.
(2) (a) He, L.; Hannon, G. J. *Nat. Rev. Genet.* **2004**, *5*, 522–531. (b) Rich, A.; RajBhandary, U. L. *Annu. Rev. Biochem.* **1976**, *45*, 805–860. (c) Montange, R. K.; Batey, R. T. *Annu. Rev. Biophys.* **2008**, *37*, 117–133.

(3) (a) Duncan, C. D.; Weeks, K. M. *Biochemistry* **2008**, *47*, 8504–8513. (b) Watts, J. M.; Dang, K. K.; Gorelick, R. J.; Leonard, C. W.; Bess, J. W., Jr.; Swanstrom, R.; Burch, C. L.; Weeks, K. M. *Nature* **2009**, *460*, 711–716.
(4) Wilkinson, K. A.; Vasa, S. M.; Deigan, K. E.; Mortimer, S. A.; Giddings, M. C.; Weeks, K. M. *RNA* **2009**, *15*, 1314–1321.
(5) (a) Merino, E. J.; Wilkinson, K. A.; Coughlan, J. L.; Weeks, K. M. *J. Am. Chem. Soc.* **2005**, *127*, 4223–4231. (b) Gherghe, C. M.; Shajani, Z.; Wilkinson, K. A.; Varani, G.; Weeks, K. M. *J. Am. Chem. Soc.* **2008**, *130*, 12244–12245.
(6) Wilkinson, K. A.; Merino, E. J.; Weeks, K. M. *Nat. Protoc.* **2006**, *1*, 1610–1616.
(7) (a) Vincent, H. A.; Deutscher, M. P. *J. Biol. Chem.* **2006**, *281*, 29769–29775. (b) Vincent, H. A.; Deutscher, M. P. *J. Mol. Biol.* **2009**, *387*, 570–583. (c) Matos, R. G.; Barbas, A.; Arraiano, C. M. *Biochem. J.* **2009**, *423*, 291–301.
(8) Lalonde, M. S.; Zuo, Y.; Zhang, J.; Gong, X.; Wu, S.; Malhotra, A.; Li, Z. *RNA* **2007**, *13*, 1957–1968.
(9) Serganov, A.; Polonskaia, A.; Phan, A. T.; Breaker, R. R.; Patel, D. J. *Nature* **2006**, *441*, 1167–1171.
(10) Mortimer, S. A.; Weeks, K. M. *J. Am. Chem. Soc.* **2007**, *129*, 4144–4145.
(11) Heidenreich, O.; Pieken, W.; Eckstein, F. *FASEB J.* **1993**, *7*, 90–96.
(12) (a) Wu, S.; Skolnick, J.; Zhang, Y. *BMC Biol.* **2007**, *5*, 17. (b) Zhang, Y. *BMC Bioinformatics* **2008**, *9*, 40.
(13) (a) Frazao, C.; McVey, C. E.; Amblar, M.; Barbas, A.; Vonrhein, C.; Arraiano, C. M.; Carrondo, M. A. *Nature* **2006**, *443*, 110–114. (b) Lorentzen, E.; Basquin, J.; Tomecki, R.; Dziembowski, A.; Conti, E. *Mol. Cell* **2008**, *29*, 717–728.
(14) Barbas, A.; Matos, R. G.; Amblar, M.; López-Viñas, E.; Gomez-Puertas, P.; Arraiano, C. M. *J. Biol. Chem.* **2009**, *284*, 20486–20498.
(15) (a) Shapiro, R.; Hachmann, J. *Biochemistry* **1966**, *5*, 2799–2807. (b) Ehresmann, C.; Baudin, F.; Mougel, M.; Romby, P.; Ebel, J. P.; Ehresmann, B. *Nucleic Acids Res.* **1987**, *15*, 9109–9128.
(16) (a) Edwards, T. E.; Ferre-D'Amare, A. R. *Structure* **2006**, *14*, 1459–1468. (b) Lang, K.; Rieder, R.; Micura, R. *Nucleic Acids Res.* **2007**, *35*, 5370–5378. (c) Rentmeister, A.; Mayer, G.; Kuhn, N.; Famulok, M. *Nucleic Acids Res.* **2007**, *35*, 3713–3722. (d) Kulshina, N.; Edwards, T. E.; Ferre-D'Amare, A. R. *RNA* **2010**, *16*, 186–196. (e) Ali, M.; Lipfert, J.; Seifert, S.; Herschlag, D.; Doniach, S. *J. Mol. Biol.* **2010**, *396*, 153–165.
(17) Deigan, K. E.; Li, T. W.; Mathews, D. H.; Weeks, K. M. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 97–102.

JA103781U